Analyzing Relationships Among Features of Diabetes-Induced Cerebral Microvascular Disease Using Causal Inference Methods

Lauren Shen

Morgantown High School

Diabetes and CVD: The Issue

- Type 2 diabetes mellitus (T2DM) affects 462 million people worldwide [1].
- T2DM can cause cerebral microvascular disease (CVD), affecting the brain.
- Exact relationship among features of CVD remains unclear.
  - If more was known, treatment of patients could improve.
- Machine learning (ML) in this issue is limited.
  - ML: the use of algorithms to find relationships among data

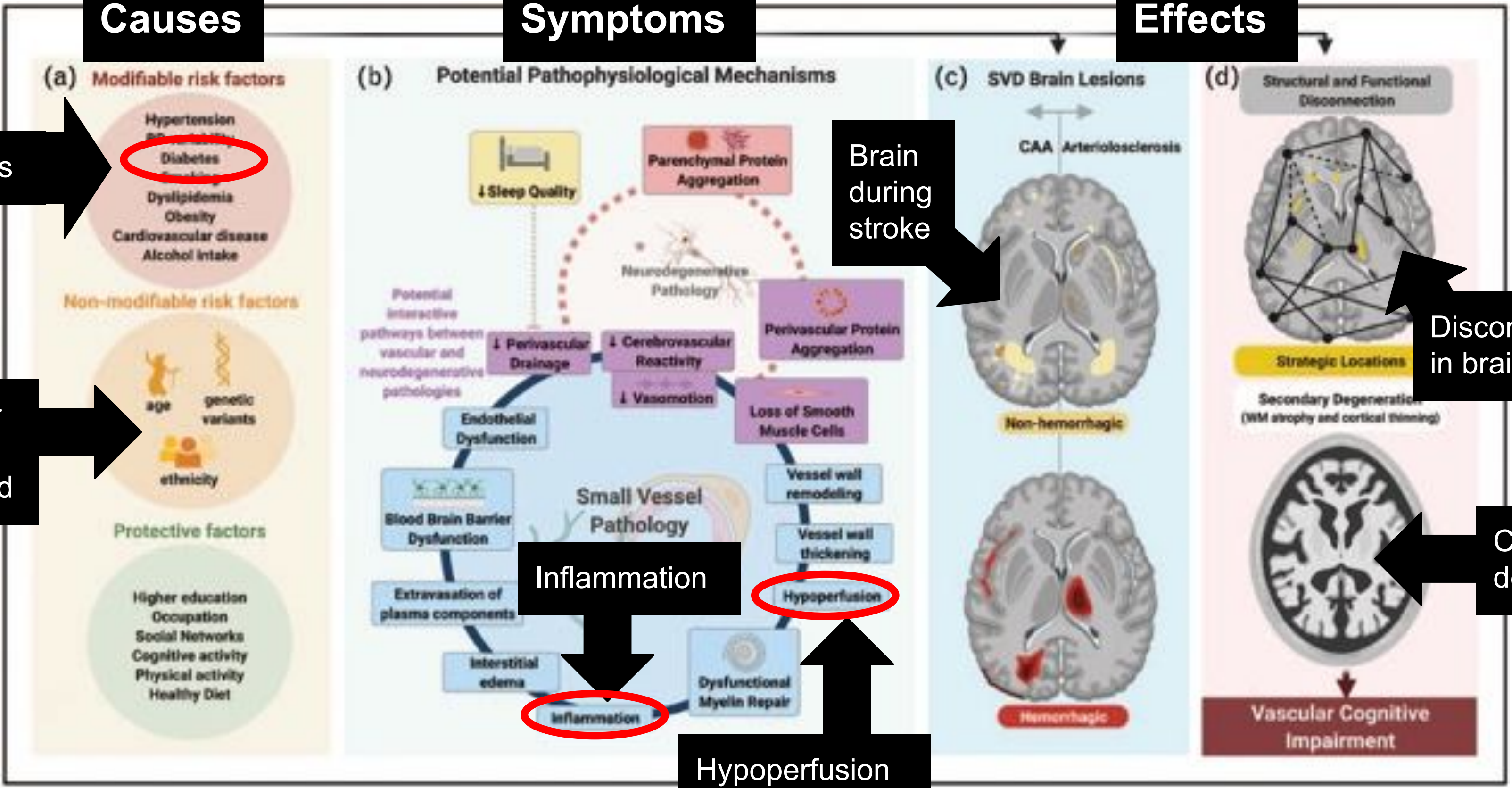
Causes, Symptoms, and Effects of CVD

[2]

Causes

Symptoms

Effects



Diabetes

Gender not included

Brain during stroke

Disconnection in brain

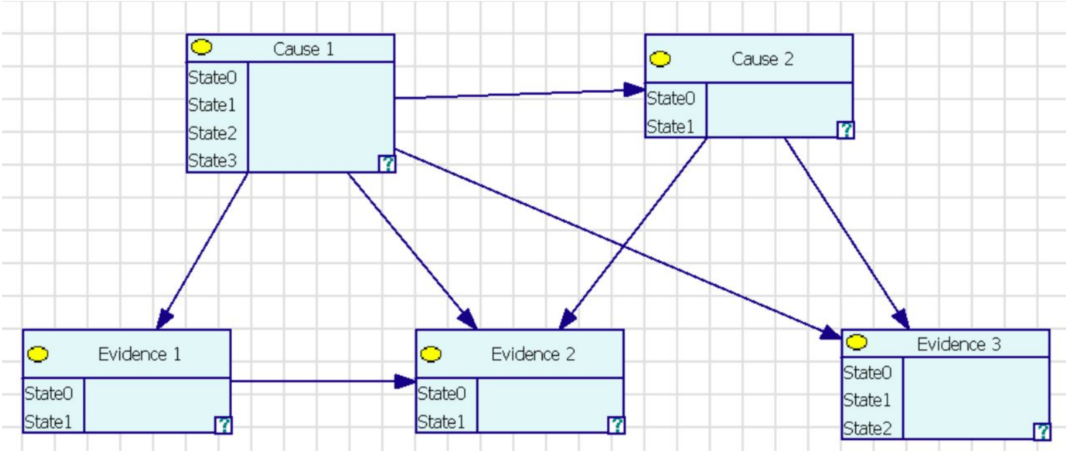
Cognitive decline

Inflammation

Hypoperfusion

Machine Learning Causal Inference in this Area

- Machine learning causal inference methods - find cause-and-effect relationships among data
- Bayesian networks - causal inference
- Previously used in predicting diabetes and stroke
- Granger causality-inspired method



[3]

Research Questions and Real World Applications

- How are severity of T2DM, gender, inflammation biomarkers, and cerebral vasoreactivity in the brain related in the context of CVD?
  - How can machine learning algorithms be applied to this issue to further explore and define the relationship between these factors?
- How can further knowledge of the relationships between the features of CVD aid patients of related diseases, such as T2DM, stroke, Alzheimer’s, and dementia?

Data Acquisition, Subsetting, and Cleaning

Distribution of Data Points in Each Subset

Subset	Male	Female	Control	T2DM	All
GE 79-1	23	49	29	43	72
GE 71, 75, 79-1	101	108	42	147	209
GE 75, 79-1	72	86	42	116	158

Data acquisition [4]

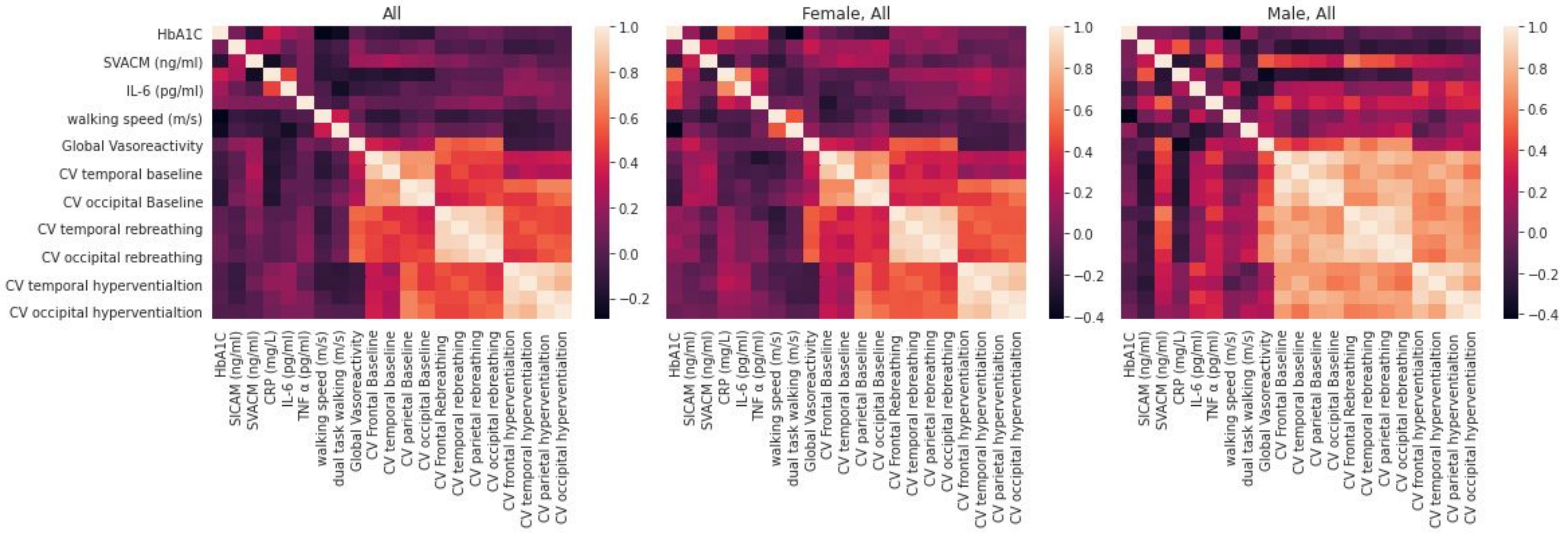
Linear Regression, Pearson’s r, Correlation Heatmaps

- Linear regression and Pearson’s r found from pairs of features.
- Correlation heatmaps visualized using data from GE 79-1.
  - Split based on group and gender.

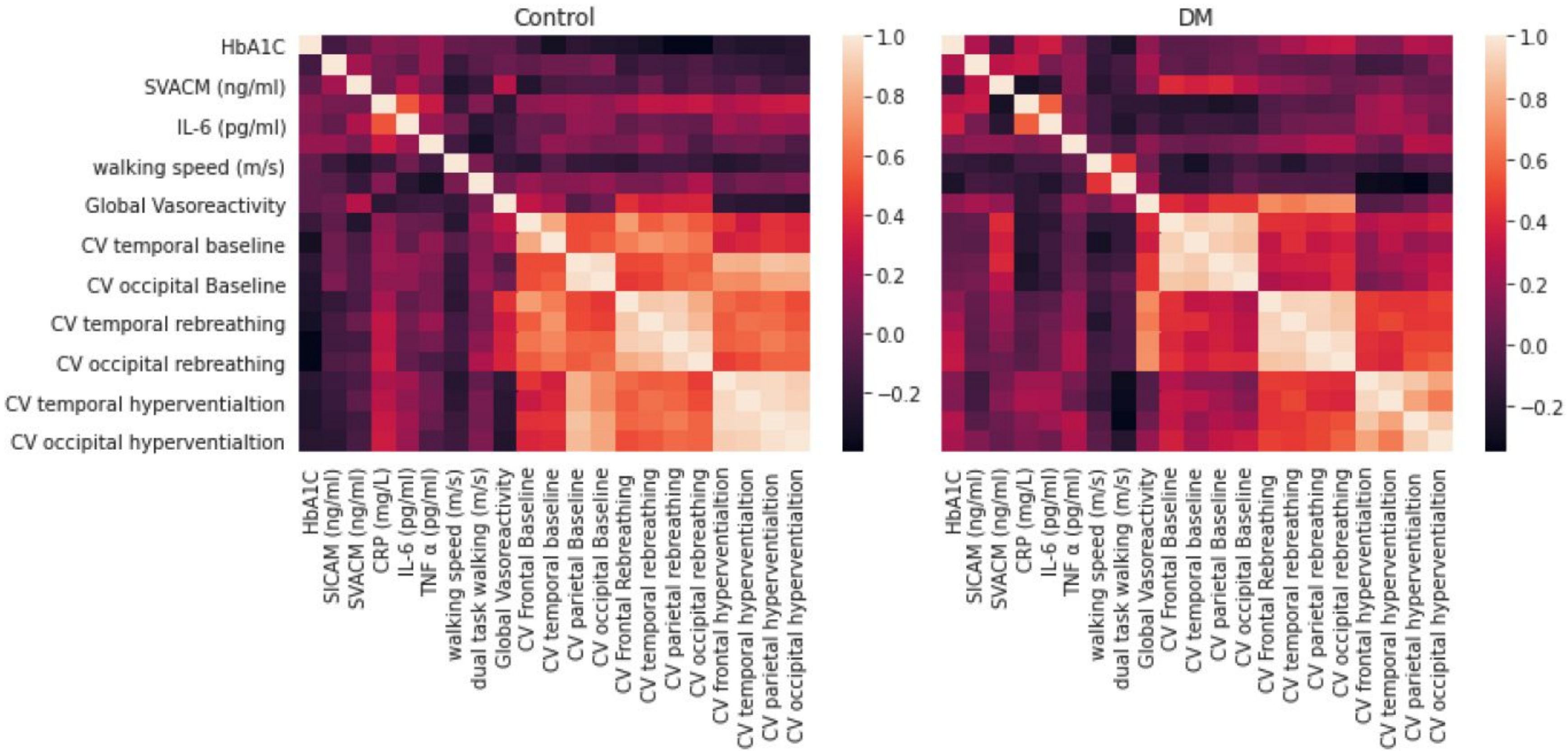
Distribution of Data Points in GE 79-1

	Male	Female	All
Control	8	21	29
T2DM	15	28	42
All	23	49	72

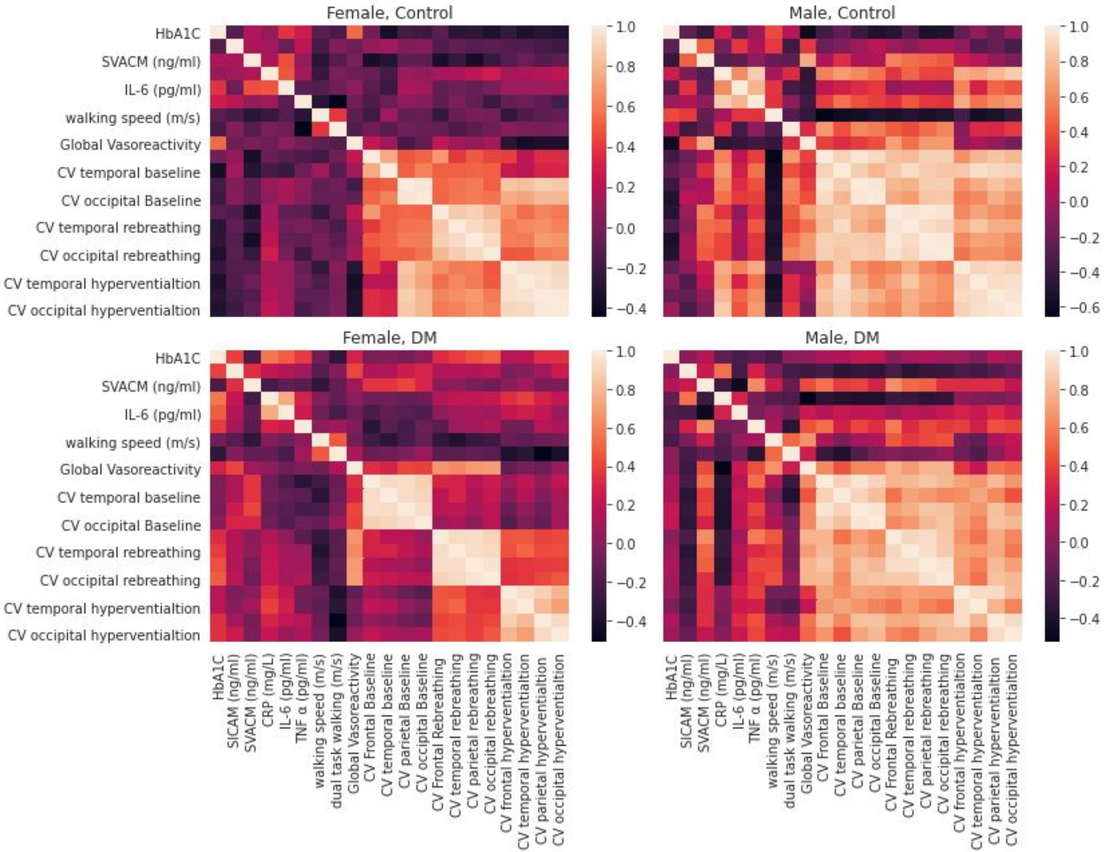
Correlation Heatmaps (All, Female, and Male)



Correlation Heatmaps (Control and DM)



Correlation Heatmaps (Gender and Group)



Why are there differences?

- Stroke has different effects in males and females [6]
- Sex-specific inflammatory signalling [5]
- SVCAM and ischemic stroke [6]

Bayesian Networks and Data Subsets

- Bayesian networks
  - The Chow-Liu algorithm was implemented on the following subsets of the data.

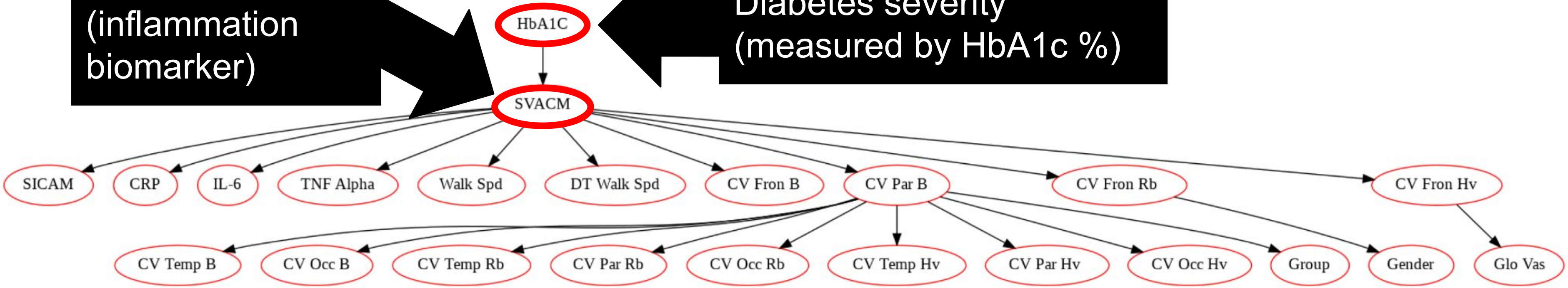
Features in Each Subset of the Data

Subset	Features
GE 79-1	Group, Gender, Severity, Inflammation Biomarkers, Global Vasoreactivity, Walking and Dual Walking Speeds, Cerebral Vasoreactivities
GE 71, 75, 79-1	Group, Gender, Severity, Walking Speed
GE 71, 75, 79-1	Group, Gender, Severity, Walking Speed, Cerebral Vasoreactivities
GE 75, 79-1	Group, Gender, Severity, Walking Speed, Global Vasoreactivity
GE 75, 79-1	Group, Gender, Severity, Inflammation Biomarkers Walking Speed Global Vasoreactivity

Bayesian Networks on Subset of Data (GE 79-1)

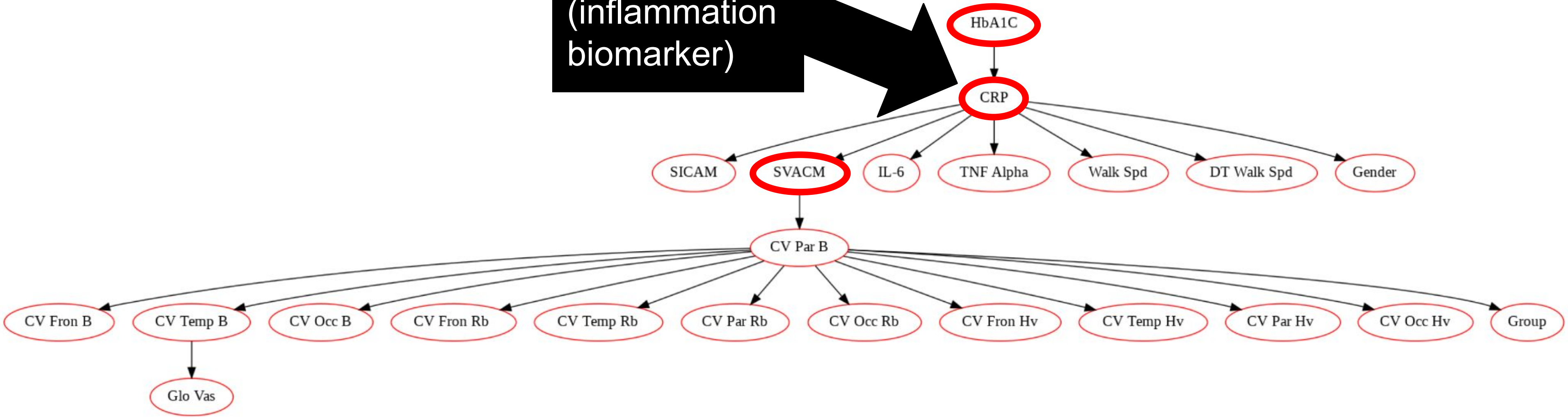
SVCAM (inflammation biomarker)

Diabetes severity (measured by HbA1c %)



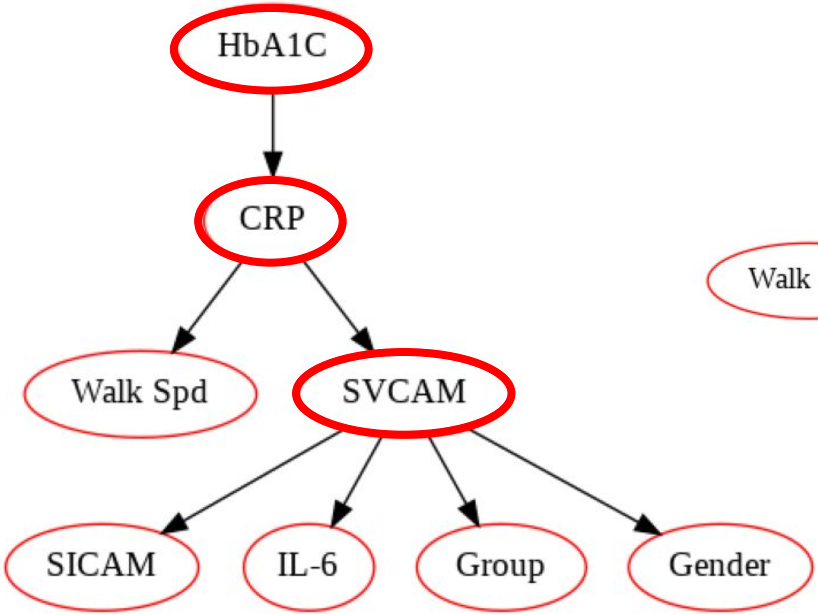
Bayesian Networks on Diabetes Group (GE 79-1)

CRP (inflammation biomarker)

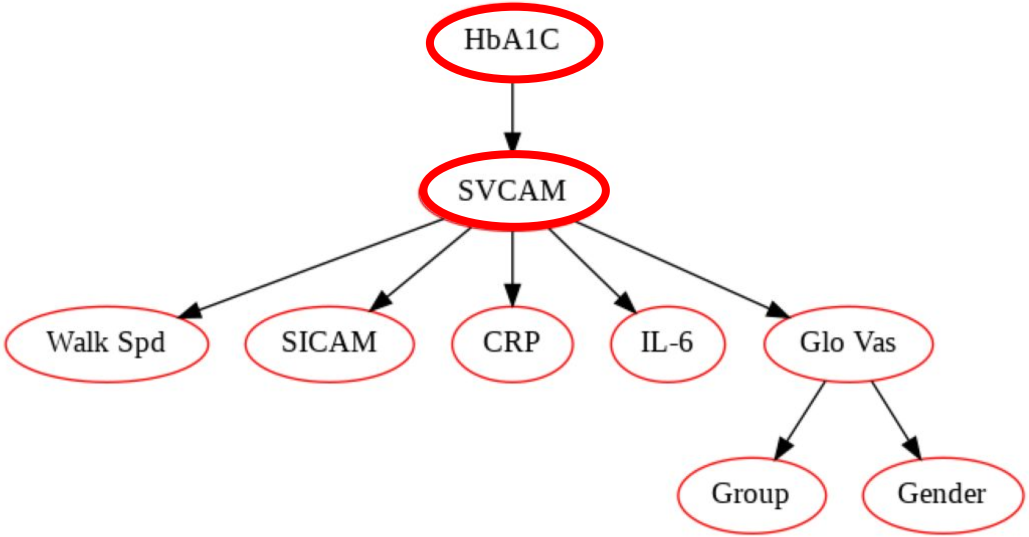


Bayesian Networks on Subsets of Data

GE 75, 79-1



GE 71, 75, 79-1



Granger Causality-Inspired Method

- Borrowing logic from Granger causality
  - In time series data, if A “Granger-causes” B, then A can be used to predict B
- Similarly:
  - We have three factors (A, B, and C) and we want to see if C “causes” A
  - Fit a regression model to see how well B can be used to predict A
  - Then, see how well B and C can be used to predict A
  - Calculate mean squared error (MSE) of each
  - Calculate ratio of AB MSE to ABC MSE
  - If ratio is significantly greater than 1, C likely has a factor in predicting A
- Linear regression and SVMs were used

Granger-Inspired Method

Linear regression:

- A: IL-6, B: CRP
- C: Severity, SVCAM, Walking Speed

SVM:

- A: Severity, B: SVCAM
- C: CRP, IL-6, Walking Speed

Results of the Granger-Inspired Method

Model	A, B, C	AB Test	ABC Test	AB/ABC MSE Ratio
Linear	A: IL-6 B: CRP C: HbA1c %	0.7248	0.3920	1.8490
Linear	A: IL-6 B: CRP C: SVCAM	0.7248	0.4756	1.5240
Linear	A: IL-6 B: CRP C: Walking Spd	0.7248	0.4780	1.5163
SVM	A: HbA1c % B: SVCAM C: CRP	2.4998	1.7742	1.4089
SVM	A: HbA1c % B: SVCAM C: IL-6	2.4498	1.7746	1.4086
SVM	A: HbA1c % B: SVCAM C: Walking Spd	2.4998	1.7752	1.4082

Granger-Inspired Method

Linear regression:

- A: IL-6, B: CRP
- C: Severity, SVCAM, Walking Speed

SVM:

- A: Severity, B: SVCAM
- C: CRP, IL-6, Walking Speed

Results of the Granger-Inspired Method

Model	A, B, C	AB Test	ABC Test	AB/ABC MSE Ratio
Linear	A: IL-6 B: CRP C: HbA1c %	0.7248	0.3920	1.8490
Linear	A: IL-6 B: CRP C: SVCAM	0.7248	0.4756	1.5240
Linear	A: IL-6 B: CRP C: Walking Spd	0.7248	0.4780	1.5163
SVM	A: HbA1c % B: SVCAM C: CRP	2.4998	1.7742	1.4089
SVM	A: HbA1c % B: SVCAM C: IL-6	2.4498	1.7746	1.4086
SVM	A: HbA1c % B: SVCAM C: Walking Spd	2.4998	1.7752	1.4082

Main Findings

- Main findings:
    - HbA1C %, SVCAM, CRP, IL-6, and walking speed seem to be the most important features among the ones we examined.
    - Differences between genders for cerebral vasoreactivities.
        - Partly explained by sex-specific inflammation biomarkers in ischemic stroke.
- Applications include:
    - Testing for HbA1C %, SVCAM, CRP, and IL-6 levels aid in finding early signs of diseases related to CVD [7].
    - Diabetes patients tested to find early signs of cognitive decline.

Future Directions

- Include different features of CVD
- Determine strength of relationship between HbA1C %, SVCAM, CRP, IL-6, and walking speed
- Work with more data
- Apply different Bayesian network algorithms and causal inference methods
- Studies in adjacent fields, such as in CVD-related stroke and CVD-related cognitive decline.

References

[1] Khan, M., Hashim, M. J., King, J. K., Govender, R. D., Mustafa, H., Al Kaabi, J. (2020). Epidemiology of Type 2 Diabetes - Global Burden of Disease and Forecasted Trends. Journal of epidemiology and global health, 10(1), 107–111. https://doi.org/10.2991/jegh.k.191028.001

[2] Zanon Zotin, M. C., L. Sveikata, A. Viswanathan & P. Yilmaz (2021) Cerebral small vessel disease and vascular cognitive impairment: from diagnosis to management. Curr Opin Neurol, 34, 246-257.

[3] Horný, M. (2014). Bayesian networks. *Boston University School of Public Health*, *17*.

[4] Novak, V., Quispe, R., & Saunders, C. (2021). Cerebral perfusion and cognitive decline in type 2 diabetes (version 1.0.0). *PhysioNet*. https://doi.org/10.13026/rbeh-9r20.

[5] Tchalla, A. E., Wellenius, G. A., Travison, T. G., Gagnon, M., Iloputaife, I., Dantoine, T., Sorond, F. A., & Lipsitz, L. A. (2015). Circulating vascular cell adhesion molecule-1 is associated with cerebral blood flow dysregulation, mobility impairment, and falls in older adults. *Hypertension (Dallas, Tex. : 1979)*, *66*(2), 340–346. https://doi.org/10.1161/HYPERTENSIONAHA.115.05180

[6] Spychala MS, Honarpisheh P, McCullough LD. Sex differences in neuroinflammation and neuroprotection in ischemic stroke. J Neurosci Res. 2017 Jan 2;95(1-2):462-471. doi: 10.1002/jnr.23962. PMID: 27870410; PMCID: PMC5217708.

[7] Landry, A., Docherty, P., Ouellette, S., & Cartier, L. J. (2017). Causes and outcomes of markedly elevated C-reactive protein levels. *Canadian family physician Medecin de famille canadien*, *63*(6), e316–e323.